

Curso Inteligência Artificial: do Zero ao Infinito

Datasets e Formatos de Anotação

Universidade Federal de Mato Grosso

Agenda

1 Introdução

- COCO Dataset
- Pascal VOC Dataset

2 Annotations Format

- COCO Format
- Pascal VOC Format
- Albumentations Format
- YOLO Format

3 Annotations Tools

4 Extra: Conversion Tools

Datasets

Os principais datasets para a tarefa de detecção de objetos são:

- COCO Dataset
- PASCAL VOC dataset

Datasets

COCO Dataset

- O **Microsoft COCO Dataset** (*Common Objects in Context*) é um conjunto de dados de detecção e segmentação de objetos publicado pela Microsoft.
- Possui 330K imagens (200K+ anotadas).
- Possui >2M instâncias em 80 classes.
- Utiliza o formato JSON para as anotações.

Datasets

PASCAL VOC Dataset

- O **PASCAL VOC Dataset** fornece um dataset de imagens padronizado para detecção de objetos.
- Possui $\approx 11K$ imagens contendo $\approx 27K$ objetos anotados pertencentes à 20 classes.
- Utiliza o formato XML para as anotações.

Annotations Format

Os principais formatos para anotação são definidos pelos dois datasets:

- **Pascal VOC Dataset**
- **COCO Dataset**

Datasets

COCO Dataset Format

O formato **COCO Dataset** utiliza um arquivo **JSON** de anotação, o qual possui os seguintes blocos:

- **info**: contém informações sobre o dataset.
- **licenses**: contém uma lista de licenças das imagens.
- **categories**: contém uma lista de categorias.
- **images**: contém informações de todas as imagens, exceto bounding boxes.
- **annotations**: contém lista dos *bounding boxes*.

Datasets

COCO Dataset

```
"info"  
{  
  "year": int,  
  "version": str,  
  "description": str,  
  "contributor": str,  
  "url": str,  
  "date_created": datetime,  
}
```

```
"info":  
{  
  "year": 2019,  
  "version": "1.0",  
  "description": "Flower and Fruits dataset",  
  "contributor": "Flowers Inc.",  
  "url": "http://test.org",  
  "date_created": "2019/12/04"  
}
```


Datasets

COCO Dataset

```
licenses
[  
  {  
    "id": int,  
    "name": str,  
    "url": str,  
  }  
]
```

```
"licenses":  
[  
  {  
    "id": 1,  
    "name": "Free License",  
    "url": "http://flower.org/",  
  }  
]
```

Datasets

COCO Dataset

```
Categories
[{"id": int,
 "name": str,
 "supercategory": str,
}]
```

```
"categories":
[
  {"id": 1, "name": "rose", "supercategory": "flower"},
  {"id": 2, "name": "tulip", "supercategory": "flower"},
  {"id": 10, "name": "Apple", "supercategory": "fruit"}
]
```

Datasets

COCO Dataset

```
image{  
  "id": int,  
  "width": int,  
  "height": int,  
  "file_name": str,  
  "license": int,  
  "flickr_url": str,  
  "coco_url": str,  
  "date_captured": datetime,  
}
```

```
"images": [  
  {  
    "id": 397133,  
    "width": 640,  
    "height": 640,  
    "file_name": "101.jpg",  
    "license": 1,  
    "date_captured": "2019-12-04 17:02:52"  
  },  
  {  
    "id": 397122,  
    "height": 640,  
    "width": 640,  
    "file_name": "102.jpg",  
    "license": 1,  
    "date_captured": "2019-12-04 17:02:52"  
  }  
]
```

Datasets

COCO Dataset

```
annotation{
  "id": int,
  "image_id": int,
  "category_id": int,
  "segmentation": RLE or [polygon],
  "area": float,
  "bbox": [x,y,width,height],
  "iscrowd": 0 or 1,
}

"annotations": [
  {
    "segmentation":
    [[510.66,423.01,511.72,420.03,...,510.45,423.01]],
    "area": 702.10,
    "iscrowd": 0,
    "image_id": 397133,
    "bbox": [433.07,355.93,138.65,228.67],
    "category_id": 18,
    "id": 1768
  },
  {
    "segmentation":
    {
      "counts": [12,56,198,10]
      "size": [120, 240]
    }
    "area": 500.2,
    "iscrowd": 1,
    "image_id": 397122,
    "bbox": [473.07,395.93,38.65,28.67],
    "category_id": 18,
    "id": 1768
  }
]
```

Datasets

Pascal VOC Dataset

O formato **Pascal VOC Dataset** utiliza um arquivo **XML** de anotação, o qual possui os seguintes blocos:

- **Folder**: nome da pasta que contém o arquivo.
- **Filename**: nome do arquivo existente na pasta.
- **Size**: informações sobre as dimensões da imagem.
- **Object**: informações sobre o objeto a ser identificado.

Datasets

Pascal VOC Dataset

As informações sobre o objetos são:

- **Name:** nome do objeto a ser identificado.
- **Truncated:** informa se o objeto é visível parcialmente na imagem.
- **Difficult:** indica se o objeto é difícil de detectar.
- **Bounding box:** coordenadas do *bounding box*.

Datasets

Pascal VOC Dataset

```
<annotation>
  <folder>Kangaroo</folder>
  <filename>00001.jpg</filename>
  <path>./Kangaroo/stock-12.jpg</path>
  <source>
    <database>Kangaroo</database>
  </source>
  <size>
    <width>450</width>
    <height>319</height>
    <depth>3</depth>
  </size>
  <segmented>0</segmented>
  <object>
    <name>kangaroo</name>
    <pose>Unspecified</pose>
    <truncated>0</truncated>
    <difficult>0</difficult>
    <bndbox>
      <xmin>233</xmin>
      <ymin>89</ymin>
      <xmax>386</xmax>
      <ymax>262</ymax>
    </bndbox>
  </object>
</annotation>
```

Datasets

Pascal VOC vs COCO Dataset Format

As principais diferenças entre os formatos são:

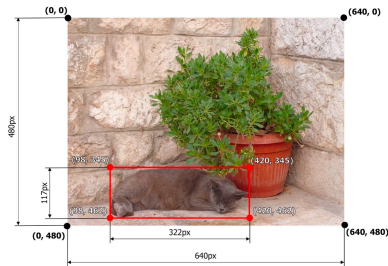
- Pascal VOC usa anotações no formato XML, um arquivo XML por imagem.
- COCO Dataset utiliza um único arquivo JSON para armazenar as anotações.

Albumentations Format

- O formato Albumentations é similar ao Pascal VOC.
- Difere em relação às coordenadas do bounding boxes, pois utiliza valores normalizados.
- Para normalizar os valores, divide-se as coordenadas em pixels dos eixos x e y pela altura e largura da imagem.

Fonte: https://albumentations.ai/docs/getting_started/bounding_boxes_augmentation/

Bounding Boxes Format



pascal_voc

[x_min, y_min, x_max, y_max] → [98, 345, 420, 462]

albumentations

normalized [x_min, y_min, x_max, y_max] → [0.153125, 0.71875, 0.65625, 0.9625]

coco

[x_min, y_min, width, height] → [98, 345, 322, 117]

yolo

normalized [x_center, y_center, width, height] → [0.4046875, 0.8614583, 0.503125, 0.24375]

YOLO Format

- O formato YOLO utiliza um arquivo TXT.
 - ▶ YOLO refere-se ao modelo de mesmo nome.
 - Cada arquivo .TXT possui anotações referentes a uma imagem.
 - As anotações possuem o seguinte formato:
- <object-class> <x> <y> <width> <height>
- Para cada objeto presente na imagem, um linha é criada.

0 45 55 29 67

1 99 83 28 44

Annotations Tools

Algumas ferramentas gratuitas para anotação:

- 1 MakeSense.AI
- 2 LabelImg
- 3 VGG image annotator
- 4 LabelMe
- 5 Scalable
- 6 RectLabel

Referencias

- COCO and Pascal VOC data format for Object detection
 - ▶ <https://towardsdatascience.com/coco-data-format-for-object-detection-a4c5eaf518c5>
- Image Data Labelling and Annotation ? Everything you need to know
 - ▶ <https://towardsdatascience.com/image-data-labelling-and-annotation-everything-you-need-to-know-86ede6c684b>
- Annotate Your Image - Using Online Annotation Tool!
 - ▶ <https://towardsdatascience.com/annotate-your-image-using-online-annotation-tool-52d0a742daff>
- COCO Dataset
 - ▶ <https://cocodataset.org/>
- Pascal VOC Dataset
 - ▶ <http://host.robots.ox.ac.uk/pascal/VOC/>

Referencias

- Preparing Custom Dataset for Training YOLO Object Detector
 - ▶ <https://www.visiongeek.io/2019/10/preparing-custom-dataset-for-training-yolo-object-detector.html>
- Image Data Labelling and Annotation ? Everything you need to know
 - ▶ <https://towardsdatascience.com/image-data-labelling-and-annotation-everything-you-need-to-know-86ede6c684b>

Conversion Tools

Algumas ferramentas/tutoriais para conversão entre os formatos:

- How To Convert COCO JSON to Pascal VOC XML
 - ▶ <https://roboflow.com/convert/coco-json-to-pascal-voc-xml>
- How To Convert Pascal VOC XML to COCO JSON
 - ▶ <https://roboflow.com/convert/pascal-voc-xml-to-coco-json>
- How To Convert Pascal VOC XML to YOLO Darknet TXT
 - ▶ <https://roboflow.com/convert/pascal-voc-xml-to-yolo-darknet-txt>
- How to Convert PASCAL VOC to YOLO Darknet Annotations without GitHub Code
 - ▶ <https://nanoml.org/blog/How-to-Convert-PASCAL-VOC-to-YOLO-Darknet-Annotations-without-GitHub-Code-0.php>

Curso Inteligência Artificial: do Zero ao Infinito

Datasets e Formatos de Anotação

Universidade Federal de Mato Grosso